

Neuberger Tilda 2021. A rövid és hosszú zöngétlen explozívák észlelése felnőtteknél és kisiskolás gyermekeknél. Beszédtudomány – Speech Science 2021. (közlésre elfogadott, lektorált változat)

A rövid és hosszú zöngétlen explozívák észlelése felnőtteknél és kisiskolás gyermekeknél

Neuberger Tilda

Nyelvtudományi Kutatóközpont

Abstract

There has been much research indicating that a specific speech sound is characterized by wide variability in timing of articulation. Despite the wide variability and the considerable overlap in the duration of singleton and geminate realisations, it is also confirmed by previous production studies that the primary acoustic attribute that distinguishes geminates from singletons is duration. In speech perception research, it is an essential question how listeners discriminate phonological categories of quantity along a continuous durational scale of the phonetic realisations. This perceptual task seems to be more difficult for children than for adults. The aim of this paper is to examine the role of duration as a perceptual distinction between single and geminate stops using a binary discrimination test of stops with systematically manipulated closure duration. The main question was what kind of similarities and differences can be found between 7-8-year-old children' and adults' perceptual shifts from singleton to geminate. Results confirmed that closure duration is a sufficient and adequate perceptual cue in the distinction of single and geminate stops both in children and adults; however, perceptual shifts differed between children and adults, and depended on the place of articulation. Moreover, displacement of the response curves based on the original quantity of stimuli (originally geminate or originally singleton) indicated that cues other than closure duration may contribute to the length contrast in Hungarian stops.

1. Bevezetés

1.1 Elméleti háttér

Általános kiindulópont a laboratóriumi fonológiai kutatásokban, hogy a fonológiailag releváns megkülönböztető jegyek (pl. zöngésség, hosszúság) a beszédhangok szintjén is realizálódnak az artikulációban, az akusztikumban és a percepcióban. A kvantitásra vonatkozó [\pm hosszú] disztinkatív jegyet az időtartamhoz, illetőleg az ejtés, hangoztatás tartamához szokás kapcsolni, eszerint megkülönböztetünk rövid (szingleton) és hosszú

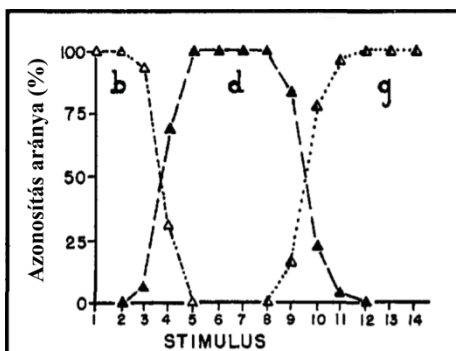
(gemináta) mássalhangzókat. A nemzetközi szakirodalomban használt terminológia szerint a tanulmányban szinonimaként használjuk a *rövid mássalhangzó* és a *singleton*, valamint a *hosszú mássalhangzó* és a *gemináta* terminusokat.

Számos akusztikai fonetikai kutatás alátámasztotta, hogy a rövid és a hosszú mássalhangzók megkülönböztetésében az elsődleges akusztikai attribútum az időtartam (lásd Ridouane, 2010 összefoglalója 24 nyelvről). Univerzális jelenségként elmondható, hogy a geminátákra a világ nyelveiben átlagosan 1,5–3-szor hosszabb időtartam jellemző a rövid konzonánsokhoz képest (Ladefoged & Maddieson, 1996). Az időtartam adatok azonban nagyfokú variabilitást mutatnak beszélőn belül és beszélők között is, amely a rövid és hosszú hangzók abszolút időtartamainak eloszlása között jelentős átfedést okoz (pl. Hirata & Whiton, 2005; Olaszy, 2006). Ennek különféle okai lehetnek a beszédhang kontextus-függő időviszonyaitól kezdve a beszélő artikulációs sajátosságaiig. Az artikulációs tempó különbségeinek és változékonyságának eredményeképpen például egy gyors artikulációs tempóban ejtett gemináta akár rövidebb időtartamú is lehet, mint egy rövid konzonáns.

Felmerül a kérdés, hogy a beszédhangok sokfélesége ellenére miként lehet az észlelésük a mindennapi kommunikációban mégis sikeres? A beszédpercepció kísérletek nagy része fókuszál annak vizsgálatára, hogy a beszédhangok különféle megvalósulásait figyelembe véve mik azok a releváns jegyek, amelyek lehetővé teszik az absztrakt fonémákhoz való hozzárendelésüket. Másképp fogalmazva: a jelentős mértékű variabilitás ellenére léteznek-e állandó, invariáns jegyek, amelyek a percepció alapjául szolgálnak? Ez a kérdés áll az invariancia-alapú beszédpercepció elméletek középpontjában, de a legtöbb beszédészlelési modell kitér rá. A következőkben megemlítünk néhányat, amelyek téziseit a jelen kutatás témájához közelebb állónak érezzük.

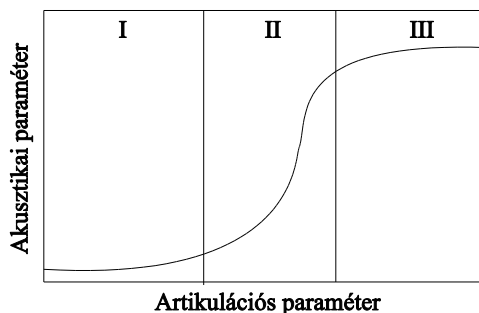
A **motoros elmélet** (motor theory: Liberman & Mattingly, 1985), valamint a **közvetlen realista elmélet** (direct realist theory: Fowler, 1986) közös pontja, hogy összeköti az artikulációt a percepcióval. Feltételezik, hogy a hallgatók nem akusztikai fonetikai információkat, hanem gesztusokat nyernek ki az észlelés során. Előbbi képviselői az észlelés invariáns alapegységét a beszélő szándékolt artikulációs gesztusaiban (intended phonetic gestures) látják, amelyeket a hallgató az akusztikai jelen keresztül a saját, produkcióról tárolt ismeretei alapján dekódol. Utóbbi középpontjában pedig az aktuális artikulációs gesztusok állnak, és nem az azokat megelőző szándékolt gesztusok, vagy az absztrakt fonémák. Vélhetően az akusztikai jelben található információ elegendő a gesztus felismeréséhez. Liberman és munkatársai (1957) azon kísérletük alapján, amelyben a hallgatók egy

kontinuum mentén folyamatosan változó szintetizált hangokat (/ba/ - /da/ - /ga/) diszkrét kategóriákként azonosítottak, felvetették a kategorikus észlelés fogalmát (1. ábra).



1. ábra: Kategoriai azonosítás (Liberman et al., 1957: 361)

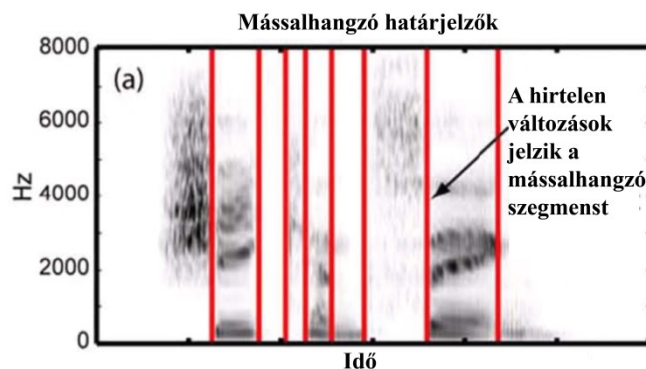
A **kvantális elmélet** (quantal theory: Stevens, 1989) az akusztikai jelben keresi az invariáns jegyeket. Kimondja, hogy az artikuláció és az akusztikum, valamint az akusztikum és a percepció közötti kapcsolat nem lineáris, hanem ugrásszerű (2. ábra). Ezt bizonyítja, hogy az akusztikai szerkezetben tapasztalható relatíve nagy eltérés esetén a hallgató nem feltétlenül észlel különbséget, ugyanakkor a percepció különbségtételben olykor apró akusztikai eltérések is szerepet játszanak. Az észlelést kategoriálisnak tartja. Az akusztikai paraméterek közül az időtartam kihívást jelent a kvantális interpretáció számára: nem tisztázott, hogy a szegmensek fizikai időtartamának széles változatosságát tekintve a hallgatók miként használják fel az idő-alapú információt a rövid és hosszú jegyek megkülönböztetéséhez.



2. ábra: Az artikuláció és az akusztikum kapcsolata (Stevens, 1989: 4)

Stevens (2002) későbbi modelljében hangsúlyozza a **megkülönböztető jegyek** (mint a fonológiai kontraszt alapegységeinek) szerepét, amelyekhez akusztikai kulcsok kapcsolhatók, és amelyek mentén a hallgató bináris döntéseket hoz a percepció során (vagyis az észlelés alapegységeinek tekinti). A megkülönböztető jegyek (pl. [mássalhangzó], [szonoráns]) a legkönnyebben az akusztikai jel azon részén azonosíthatók, ahol ún. **akusztikai határjelző**

(landmark) található. Ezek olyan régiói a beszédnek, ahol hirtelen változás jelenik meg a beszéd spektrumában, mint például egy felpattanó zárhang esetében a felpattanás pillanata (3. ábra). A hallgató a határjelzők közvetlen környezetében elemzi az akusztikai paramétereket, melynek során eljut a megfelelő megkülönböztető jegyekhez, illetve az ezek kötegéből felépülő szegmensekhez (fonémák), majd megtalálja az ezek sorozatából álló szavakat a lexikonban.



3. ábra: Határjelzők az akusztikai jelben (Stevens, 2002 alapján Slifka 2006: 95)

Az **akusztikai invariancia elmélete** (acoustic invariance theory: Blumstein & Stevens, 1981) szerint az állandó, invariáns jegyek közvetlenül az akusztikai jelben találhatóak. Az észleléshez szükséges információt a hallgatók jellemződetektáló mechanizmusok által nyerik ki a beszédjelből. Vizsgálták többek közt az explozívák képzési helyét, s kimutatták, hogy a zárfeloldás rövid idejű spektruma tartalmazza a képzési helyüket meghatározó invariáns jegyeket. Ezt olyan percepciók kísérletek eredményeivel támasztották alá, melyekben (szintetizált) CV hangsorokban a feloldás és a magánhangzó-átmenet akusztikai paramétereit szisztematikusan manipulálták, majd a zárhang azonosítását kérték a hallgatóktól (pl. Stevens & Blumstein, 1978). Az eredmények szerint a felpattanás utáni 10-20 ms-os szakasz jellemzőiben mutatkozik a zárhangok képzéshely szerinti különbsége. Mások a zöngés-zöngétlen megkülönböztetés invariáns jegyeit kívánták felderíteni a zárhangok produkciós és percepciók elemzésével (pl. Stevens & Klatt, 1974). Továbbá a rövid-hosszú megkülönböztetés akusztikai és felismerési kulcsait a tanulmányok jelentős hányada szintén a felpattanó zárhangok osztályán keresztül vizsgálta (pl. Lisker, 1958; Hankamer et al., 1989; Pickett et al., 1999; Tsukada et al., 2018; Zhou et al., 2021). Összetett akusztikai szerkezetük különösképpen indokoltá teszi a kérdés vizsgálatát az explozívák (illetve az affrikáták) körében: felvetődik ugyanis a kérdés, hogy a kvantitás a mássalhangzó komplex időszerkezetének mely elemében manifesztálódik. A munkák egy része a zárszakasz abszolút

időtartamát vette górcső alá, mások a relatív időtartamokat, vagyis a különböző paraméterek egymáshoz viszonyított szerepét hangsúlyozzák az invariáns jegyek feltérképezésében.

A zárszakasz abszolút időtartama elegendő felismerési kulcsnak bizonyult azokban a percepciók tesztekben, amelyek a záridőtartam mesterséges manipulációjával létrehozott stimulusokat alkalmaztak. Ilyen teszteket a világ számos nyelvén elvégeztek, többek között a maráthi nyelvben (Lisker, 1958), az arabban (Obrecht, 1965), a törökben és a bengáliban (Hankamer et al., 1989), a japánban és a finnben (Yoshida et al., 2015; Zhou et al., 2021) vagy az olaszban (Esposito & Benedetto, 1999). Ezen kutatások közös célja az volt, hogy megtalálják, hol található az a pont, amikor a hallgatók észlelése a rövid hang megítéléséből átvált a hosszú hang ítéletébe, vagyis hogy hol található a percepciók határ a szingleton-gemináta kategóriák között. Az eredményeik szerint a szingleton és a gemináta közötti percepciók határvonal a zöngétlen zárhangok esetén valahol a 120–180 ms közötti tartományba esik (pl. Esposito & Benedetto, 1999; Yoshida et al., 2015; Zhou et al., 2021).

A **relációs akusztikai invariancia** (relational acoustic invariance) elmélete szerint az invariáns jegyek nem önmagukban találhatók meg az akusztikai szerkezetben, hanem a jellemzők egymással való kapcsolatában (pl. Sussman et al., 1991; Amano & Hirata, 2010, 2015). A mássalhangzó-hosszúság tekintetében ez azt jelenti, hogy a relatív időtartamok, időviszonyok döntő szerepet játszanak a kategóriák észlelésében. Más szóval: nem csak (és nem elsődlegesen) a zárszakasz abszolút időtartamában érdemes a hosszúság akusztikai kulcsát keresni, hanem az időtartamok (pl. a zárszakasz és a megelőző magánhangzó időtartamának) egymáshoz viszonyított arányában. Továbbá nem hagyható figyelmen kívül az artikulációs tempó variabilitása sem. Amano & Hirata (2010) alátámasztotta azt, hogy míg a rövid-hosszú kategóriák közötti percepciók határhoz kapcsolódó záridőtartamok a tempótól függően változatos képet mutatnak (34–213 ms, átlag: 110 ms), addig a záridőtartamnak a szóidőtartamhoz viszonyított aránya egy állandó (invariáns) paraméternek tűnik a két fonológiai kategória elkülönítésében.

Az **anyanyelvi mágnes elmélete** (native language magnet theory; Kuhl et al., 2008) kimondja, hogy a gyermekek a születés utáni korai szakaszban (10 hónapos korukig) még képesek a világ nyelveiben előforduló legtöbb fonéma megkülönböztetésére, ami lehetővé teszi, hogy bármely nyelv fonémakészletét elsajátítsák, ám ezután átrendeződik a fonémakategória-rendszerük, amelyben a saját környezetük nyelve lesz az irányadó. Ettől fogva észlelésük csak az anyanyelvük fonémái közötti különbségekre lesz érzékeny. Mindez a prototípusok elvén alapul; a mindaddig hallott számos beszédhang-példányból kialakítják az adott fonémához tartozó prototípust. A prototípusok perceptuális mágnesként funkcionálnak,

amelyek a bemeneti beszédjelben a hasonló akusztikai jellemzőkkel bíró szegmenseket „magukhoz vonzzák”. Így a kategórián belül csökkennek a különbségek, viszont a prototípushoz nem hasonlító hangok esetében nő a különbség, vagyis ezek eltérő kategóriába fognak kerülni. Ez, az univerzálisból a nyelvspecifikus észlelésbe történő váltás a későbbiekben nehézségeket okozhat az idegen nyelv tanulásában, különösképpen a L2 fonológiai kategóriák elsajátításban. A mássalhangzó-hosszúságot illetően ez a jelenség mutatkozik meg azokban a percepció tesztekben, amelyekben nem anyanyelvi beszélők hosszúságészlelését vizsgálták (pl. Wilson et al., 2005; Sonu et al., 2013; Yoshida et al., 2015; Tsukada et al., 2018). Egyrészt azt találták, hogy a nyelvtanulók javarészt az abszolút időtartamokra támaszkodnak a döntéseikben, és nem használják a kontextus-, illetőleg tempófüggetlen felismerési kulcsokat, mint ahogyan az anyanyelvi beszélők; emiatt számukra nehézséget jelent a hosszúsági kontraszt azonosítása, különösképp az artikulációs tempó változatossága esetén. Másrészt felhívták a figyelmet arra, hogy a hallgatók a kvantitásnak a saját nyelvükre jellemző nyelvspecifikus vonásait veszik alapul a megkülönböztetésben akkor is, amikor idegen nyelvű rövid és hosszú mássalhangzókról kell döntést hozniuk.

1.2 A mássalhangzó-hosszúság vizsgálata a beszédfejlődésben

Az idegen nyelv tanulása mellett az anyanyelv elsajátításában ugyancsak elengedhetetlen a releváns jegyek felismerését lehetővé tevő akusztikai kulcsok szerepe, és azoknak a redundáns jegyeknek a számbavétele, amelyek irrelevánsak, tehát ignorálhatók. Csecsemő- és gyermekkorban a bemeneti beszédjel (input) elemzésén keresztül fejlődik a percepció és a produkció mechanizmus, vagyis az akusztikai jelenségek észlelése, illetőleg az artikulációs parancsok végrehajtása. A gyermekek fonológiai fejlődésével kapcsolatban a kutatók egy része úgy tartja, hogy a fent említett folyamatokhoz minden bizonnyal szükség van invariáns jegyekre, amelyek alapján a gyermek felfedezi anyanyelvének fonémikus szerkezetét (Klatt 1979).

Az anyanyelvi fonológia elsajátításának vizsgálatában fontos kérdés, hogy a gyermek hogyan képes a hozzá intézett folyamatos beszédet szegmentálni, különálló egységekre bontani, illetve mik lehetnek a beszédelemzés, illetőleg az észlelés alapegységei. A fonológiai fejlődés leírását célzó munkákban legtöbbször két hagyományos megközelítéssel találkozhatunk, amelyek máig jelentősek: a generatív nyelvtanban általánosan alkalmazott szabály-, vagy megszorítás-alapú elemzéssel, amelyben a fonológiai szerveződés alapegysége a szegmens (pl. Gnanadesikan 2004), valamint a kognitív megközelítéssel, melyben általában az alapegység a szó mint teljes egész (pl. Vihman & Velleman, 2000).

A gyermekek beszédészlelésére és szófelismerésére vonatkozó kutatások felfedték, hogy a gyermekek különös figyelmet fordítanak a spektrum globális mintázatára, míg a felnőttek észlelése az akusztikai jel részleteire fogékonyabb. Alátámasztották azt is, hogy – a nem anyanyelvi beszélőkhöz hasonlóan – a gyermekek másképp súlyozzák a felismerési kulcsokat az észlelésükben, mint a felnőttek (Nittrouer & Lowenstein, 2009). Ezt azt jelenti, hogy a gyermekek és a felnőttek között eltérés mutatkozik abban, hogy az akusztikai jel mely tulajdonságai lesznek számukra fonetikailag informatívabbak, vagyis kiemelt jelentőségűek (száliensek), melyek segítik a hallgatót a beszédhangokat fonémaosztályokba sorolni (fonémadöntés) vagy nagyobb, jelentéssel bíró egységekbe (szófelismerés) rendezni.

A mássalhangzók kvantitásának kérdése a gyermekek beszédfejlődésében kevésbé kutatott terület a nemzetközi szakirodalomban, mint a felnőtteknél. Vihman és Velleman (2000) megállapította, hogy az első szavaikban még azok a gyermekek is ejtenek hosszú (időtartamú) mássalhangzókat, akik olyan nyelvet sajátítanak el, ahol a rövid és hosszú mássalhangzók nem állnak fonológiai oppozícióban (pl. angol, francia). Ez azonban az anyanyelv-elsajátítás egyszavas szakaszán túl már csak azokban a nyelvekben figyelhető meg, amelyekben a kvantitás fonológiailag releváns jegy (pl. finn, japán).

A nyelvek közötti vizsgálatok arra is felhívták a figyelmet, hogy a különböző anyanyelvű gyermekek nem azonos életkorban kezdik el stabilan megkülönböztetni a produkcióban a mássalhangzókat a kvantitás alapján: a finn gyermekek például gyorsabban sajátítják el a kontrasztot, mint a japánok (Aoyama, 2000; Kunnari et al., 2001). Ennek okát a bemeneti beszéd (input) különbségeiben látják a kutatók: egyrészt a kvantitáskontraszt gyakoribb, másrészt az időtartam-különbség jelentősebb a finnben, mint a japánban. A gyakran hallott kategóriák közötti kontrasztok ugyanis előbb épülnek be a nyelvspecifikus elsajátítás során, mint a ritkábban hallottak (Anderson et al., 2003).

Khattab és Al-Tamimi (2013) a libanoni arab gyermekek produkciójában U-alakú fejlődési görbét talált a mássalhangzó-hosszúság oppozíciójában, mely szerint a kontraszt pontos realizációja egy korai fejlődési stádiumban már jelen van, majd egy visszaesés tapasztalható, és csak később alakul ki a felnőtt ejtésbeli fonológiai struktúra. Hasonlót mutattak ki a norvég szonoránsok és zöngétlen explozívák vizsgálatában a 2 és fél, a 4 és a 6 évesek produkciójában (Payne et al., 2017).

A percepció oldaláról csekély számú nemzetközi tanulmány foglalkozott a geminátákkal gyermekeknél. Egy japán csecsemőkkel vizuális habituáció-diszhabituáció (ingerre irányuló figyelem, nézéspreferencia) technikát alkalmazó kutatás (Sato et al., 2012) arról számol be, hogy a gyermekek 4 hónapos korukban még nem, de 9 és fél hónapos korukban már képesek

szingleton-gemináta diszkriminációra a zöngétlen explozívák körében, és akkor is csak a természetes ejtésű (vagyis nem a mesterségesen manipulált záridőtartamú) stimulus esetén. Ez utóbbi eredmény a természetes beszédben megjelenő kiegészítő kulcsok szerepére hívja fel a figyelmet, amelyre úgy tűnik, a kisgyermek is szenzitív. A tanulmány eredményei arra engednek következtetni, hogy ilyen fiatal korban már érzékenyvé válnak a gyermekek a kontraszt észlelésére, azonban ekkor még nem beszélhetünk fonológiai tudatosságról. A szavak belső szerkezetéhez való tudatos hozzáférés, illetve a szavak eltérő méretű egységeire történő bontás képessége későbbi életkorban (3-4 éves kortól) alakul ki, és a fonológiai tudatosság egyes területeinek (pl. rímképzés, szótagszintézis) fejlődése nyelvspecifikus sajátosságokat mutathat (Jordanidisz 2015).

1.3 A mássalhangzó-hosszúság vizsgálata magyar nyelven

Az utóbbi évtizedben egyre fokozódó érdeklődés övezi a magyar gemináták produkciós sajátosságait (pl. Siptár & Grácz 2014; Neuberger 2015; Neuberger & Beke 2017; Deme et al. 2019). A tanulmányok elsősorban a komplex akusztikai szerkezetű felpattanó zárhangokat elemezték felnőttek ejtésében. A mássalhangzó-hosszúság észlelésére vonatkozó elemzések kisebb számban készültek, és főként tágabb beszédfejlődési vizsgálatok részeként jelentek meg, amelyek célja a gyermekek fonématudatosságának feltérképezése volt (pl. Gósy, 2006; Gyarmathy & Horváth 2010; Jordanidisz 2015). Közös eredményeik alapján elmondható, hogy a gyermekkor egy bizonyos szakaszáig nehézséget jelent a kvantitáskategóriák tudatos megkülönböztetése. A fonológiai tudatosság fejlődése során a szenzitivitás és a manipulációs készség fokozatosan fejlődik a nagyobb egységektől (szótagok) a kisebbek (fonémák) felismeréséig. Továbbá a gyermekek a nagyobb különbségektől a kisebbekig haladnak az észlelés fejlődésében: a beszédhangok differenciálása során a legkönnyebb a képzés módjának eltérését felismerniük, ezt követi a képzés helye szerinti különbség, majd a zöngés-zöngétlen eltérés észlelése. Legnehezebb a hangok időtartam-különbségének felismerése.

A fenti sajátosságokat veszik figyelembe a beszédészlelést vizsgáló eljárások is. Például a 4 és 8 éves kor közötti gyermekek számára készült DIFER beszédhanghalló (fonémaészlelési) készséget vizsgáló diagnosztikus teszt (Nagy et al. 2004) arra épül, hogy a hangok differenciálása a megkülönböztető jegyek számának csökkenésével párhuzamosan nehezedik. A beszédhanghallás készsége akkor mondható kialakultnak, ha a gyermekek a csupán egy jegyben történő eltérést is érzékelni tudják.

A sztenderdizált GMP-diagnosztika a 3–13 évesek beszédfeldolgozási folyamatait méri fel (Gósy 1995/2006). A beszédhang-differenciálást mérő altesztben a gyermekek feladata annak

megítélése, hogy két elhangzott hangsor egyforma (pl. *begi-begi*) vagy különböző (pl. *voka-vokka*, *ípi-íppi*, *tegő-teggő*). Nagy mintán (hatszáz 4 és 9 év közötti gyermek) vizsgálva Gósy és Horváth (2006) azt találták, hogy a beszédhang-differenciálásban jelentős változás hét- és nyolcéves kor között tapasztalható. Mind a produkciós, mind a percepciós fejlődésben általában az utolsó lépcsőfok a fonológiai hosszúsági oppozíció megkülönböztetése; ennek stabilizálódását 7 éves korra szokás tenni (Gósy, 2006). Gyarmathy és Horváth (2010) a 6–10 év közötti, hallászavarral küzdő óvodások és kisiskolások beszédhang-megkülönböztetési képességét vette górcső alá. A képzési jegyek közül a mássalhangzó hosszúsága szerinti differenciálás jelentette a legnagyobb nehézséget a gyermekek számára, sőt még a tízéveseknek is csupán a fele tudta elkülöníteni a hangsorokat a konzonánsok időtartama szerint.

Jordanidis (2015) 4 és 10 év közötti magyar anyanyelvű gyermekek (N = 390) fonológiai tudatosságát elemezte. A fonématudatosságot mérő tesztek alapján kimutatta, hogy 6 éves kor után a gyermekek teljesítménye ugrásszerűen nő (feltehetően az olvasástanulás hatására). Az egyik altesztben a gyermekek feladata a hosszú beszédhangok azonosítása volt egy-egy adott szóban. Az eredmények szerint 6 éves korban 45%-os, 7 éves korban 73%-os, 8 éves korban pedig már 93%-os átlagteljesítményt nyújtottak a gyermekek. A hétévesek eredményei szignifikánsan eltértek a nyolcévesekétől ebben a feladatban. A szerző a beszédhangok temporális jegyeinek pontos észlelését 8 éves korra (2. osztályos korra) teszi, bár hozzáteszi, hogy eredményei alapján sokan még 3. osztályban sem észlelik a beszédhanghosszúságot megfelelően.

1.4 A jelen kutatás célja

A jelen kutatás arra keresi a választ, hogy miként különböztetik meg a rövid és a hosszú mássalhangzókat az észlelésben az időparaméter mentén a magyar anyanyelvű felnőttek és 7-8 éves kisiskolás gyermekek. A zöngétlen explozívokon végzett kutatásunk hipotéziseit az alábbiakban fogalmaztuk meg: 1. A zárképzés időtartama elégséges kulcs a kvantitásra vonatkozó döntés meghozatalában mind a felnőttek, mind a gyermekek számára. 2. A gyermekek kvantitásra vonatkozó döntései kevésbé kategorikus észlelési mintázatot (laposabb válaszgörbét) mutatnak, mint a felnőtteké, valamint 3. a rövid és hosszú hangok közötti határ a felnőttekétől eltérő időtartamértéknél húzódik az észlelésükben. 4. Továbbá feltételezzük, hogy a záridőtartamukban manipulált zárhangok eredeti kvantitása (szingletonból vs. geminátából módosított hang) befolyásolja a hallgatók ítéleteit (az akusztikai jelben található

esetleges másodlagos kulcsok hatására, amelyeket nem módosítottunk a jelen kísérletben) – függetlenül az életkortól vagy a mássalhangzók képzési helyétől.

2. Anyag, módszer, kísérleti személyek

A jelen kutatásban bináris döntési teszt segítségével vizsgáljuk a különböző életkorú (gyermek és felnőtt) hallgatók mássalhangzó-hosszúságra vonatkozó perцепciós ítéleteit. A kísérleti anyagot a Nyelvtudományi Intézet Fonetikai osztályának csendesített szobájában rögzítettük egy magyar anyanyelvű, 27 éves női beszélőtől, aki gyakorlott fonetikus lévén a hangsorok felolvasása alatt végig azonos szinten tartotta az artikulációs tempót, a dallammenetet és az átlagos intenzitást. A felvétel AT 4040 kardioid kondenzátor mikrofonnal történt a GoldWave programban, 44,1 kHz mintavételezési frekvencián, 16 bit-es tárolással. A teszthez felhasználandó anyag meghatározásához egy előkísérletet végeztünk, melynek eredményeit Neuberger (2016) ismerteti. Ennek eredményei alapján választottuk ki a jelen kísérlet anyagát: olyan VCV hangszekvenciákat (logatomokat), amelyekben az intervokális helyzetben lévő mássalhangzó szingleton vagy gemináta [p, t, k], a környező magánhangzó pedig minden esetben az [i]. Az előkísérlet résztvevői az eredeti *ipi*, *ippi*, stb. hangszekvenciákban szereplő mássalhangzókat 100%-os egyetértéssel a megfelelő kategóriaként azonosították.

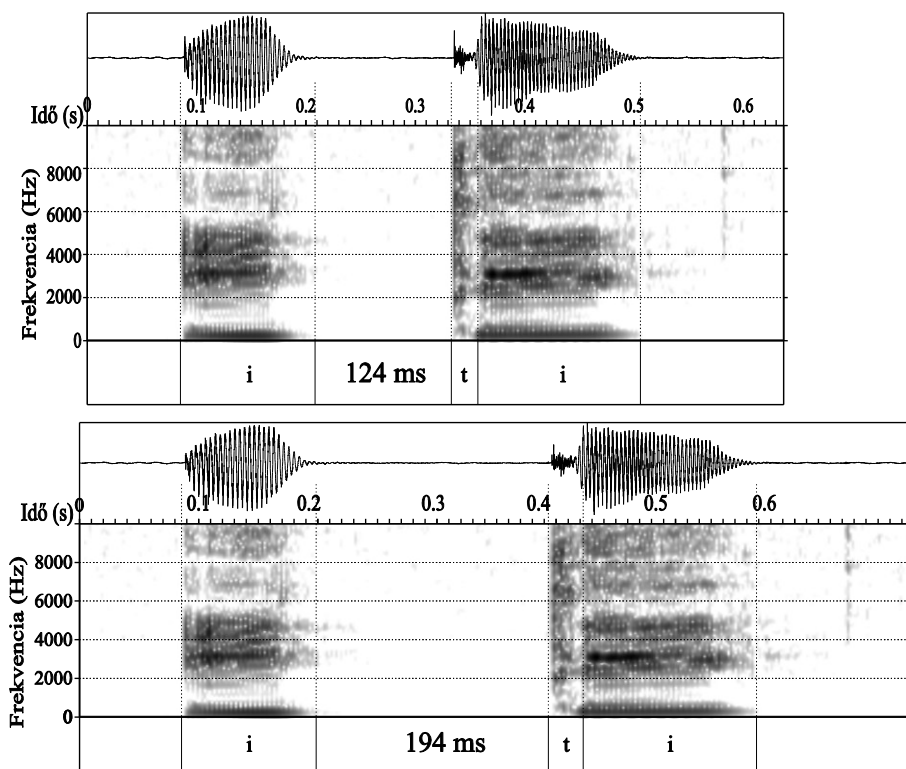
A hat kiválasztott tokenből (3 képzési hely \times 2 hosszúsági kategória) a zárszakasz szisztematikus időtartam-manipulációjával 61 stimulust hoztunk létre a következőképpen: a rövid explozívák néma szakaszát 10 ms-onként növeltük, míg elérte a hosszú pár időtartamát (1. táblázat). A kezdő rövid [p] hang például 123 ms-os volt, ezt nyolc 10 ms-os lépésben növeltük, amíg elérte a hosszú [p:] hangnak megfelelő 203 ms-ot. A hosszú explozívák zárszakaszát pedig 10 ms-os lépésekben csökkentettük, míg elérte a rövid pár időtartamát. Így a két szélső érték, valamint minden köztes lépésben kétszer szerepelt egy inger: egy szingletonból és egy geminátából előállított, azonos záridőtartamú hang. Ezt követően további lépésekkel is kibővítettük az időmanipulálást: a szingletonok időtartamán túl is csökkentettük a gemináták zársidejét a 100 ms-ot megközelítő értékekig. Arra voltunk kíváncsiak, hogy a természetes ejtésű (olvasott) szingletonnál rövidebb zárszakaszok esetén vajon egyöntetű lesz-e a hallgatók ítélete a „rövid” kategória felé. Az így kialakított kísérleti anyagban jól megfigyelhető lett a 100–200 ms közötti záridő-intervallum. Bár a képzési hely szerint különbséget mutattak az eredeti zöngétlen explozívák időértékei a szakirodalmi adatoknak megfelelően (a képzési helyben hátrafelé haladva egyre rövidebbek a záridőtartamok, pl. Grácz 2013; Neuberger 2015), a csökkentést egységesen a 100 ms körüli időértékekig

folytattuk, a labiális hang esetén tíz lépésben, az alveolárisnál tizenkettő lépésben, a velárisnál pedig kilenc lépésben (innen adódik a stimulusok nem egyenlő száma az egyes képzéshelyeken). A maximum értékek pedig a felolvasott gemináták időtartamához illesztve a jelen anyagban [p] 203 ms, [t] 224 ms és [k] 194 ms-osak voltak.

1. táblázat: A kísérletben használt stimulusok adatai

Mássalhangzó	A záridőtartam minimum értéke (ms)	A záridőtartam maximum értéke (ms)	Stimulus száma (db)
[p] – [p:]	103	203	20
[t] – [t:]	104	224	23
[k] – [k:]	104	194	18

Az időtartamok manipulálása a Praat programban történt (Boersma & Weenink 2020), melynek során csupán a zöngétlen zárhangok néma fázisát növeltük; a felpattanást, a zöngékezdési időt, valamint a környező magánhangzókat változatlanul hagytuk. A 4. ábrán látható egy példa, melyben a 124 ms-os zárszakaszt 194 ms-osra növeltük, az *i*-vel jelölt magánhangzókat és a *t*-vel jelölt felpattanást és VOT-t az eredeti ejtés szerint meghagytuk.



4. ábra: Példa a záridőtartam manipulálására a Praat programban

A percepció tesztet minden hallgatóval egyénileg végeztük el a Praat szoftverben az ExperimentMFC nevű modul segítségével. A gyermekeket egy csendes iskolai teremben, a felnőtteket egy egyetemi szobában teszteltük. Ugyanazt a rögzített anyagot hallgatták meg Sennheiser HD 419 fejhallgatón keresztül. A feladatuk az volt, hogy egy rövid stimulus (pl. [ipi]) meghallgatása után döntést hozzanak arról, hogy a hangmintában szereplő mássalhangzó rövid vagy hosszú (pl. [ipi] vagy [ip:i]). Ha a hallgató észlelése alapján rövidnek ítélte a hangot, akkor a „rövid”, ha hosszúnak, akkor a „hosszú” gombra kellett kattintania (vagy a kisiskolások esetében: rámutatnia a „rövid” vagy „hosszú” kategóriát jelképező képre). A 7-8 éves gyermekekkel – életkorukból fakadóan – játékos formában, egy kerettörténet alkalmazásával végeztük el a kísérletet. Azon gyermekek számára, akik a teszt elvégzése idejében még nem tudtak (jól) olvasni, két kép is segítette a döntést. Egy rövid és egy hosszú testű tacsót ábrázoló kép közül kellett kiválasztaniuk, hogy gazdájuk aktuálisan melyiket hívja: a „rövid” nevűt (pl. [ipi]) vagy a „hosszú” nevűt (pl. [ip:i]), és erre a kutyák testalkata utalt. A kísérlet anyagát jelentő logatomokat random sorrendben játszottuk le. Előzetesen minden adatközlőnek bemutattuk a feladatot, az első néhány próbát nem számítottuk bele a kísérleti anyagba. Az előtesztelés alapján elmondható, hogy a vizsgálatához választott résztvevők (beleértve a gyermekeket is) a feladatot megértették.

Minden adatközlő magyar anyanyelvű, egynyelvű beszélő volt. A vizsgálatához eredetileg felkért 40 gyermek között többen voltak, akik egy idő után szisztematikusan csak rövid vagy csak hosszú válaszokat adtak a tesztelés során, ezért az ő válaszaikat kizártuk az elemzésből, mert kétséges, hogy a válaszok valóban a gyermekek percepció működését tükrözték, vagy a figyelem fenntartása a feladat végéig túlzottan nehéznek bizonyult számukra. A kísérleti személyek két csoportját így 32 gyermek és 44 fiatal felnőtt alkotta. Náluk nem volt tapasztalható figyelemfenntartási nehézségből fakadó véletlenszerű jelölés.

Az összes gyermek ép halló és ép értelmű, tipikus fejlődésűnek mondható, beszédhibájuk nem volt, közöttük 16 lány és 16 fiú szerepelt. A gyermekek 7 és 8 évesek (átlagos életkor: 7,5 év); egy budapesti általános iskola első és második osztályos tanulói voltak. A vizsgálatot mindkét osztályban a tanév második félévében végeztük el. A korábbi szakirodalmi adatok alapján (lásd Bevezetés) választottuk ezt az életkort vizsgálatunk alapjául. Alátámasztották ugyanis azt, hogy a mássalhangzók hosszúságának észlelésében kiemelt jelentőségű e korosztály: az anyanyelv-elsajátítás ebben a szakaszában jelentős változás történik a kvantitásoppozíció észlelésében a korábbi életévekhez képest, illetve 7-8 éves korra elvárható a beszédhang-hosszúság szerinti diszkrimináció pontos működése.

A kísérletben részt vett felnőttek 18–27 év közöttiek (átlagos életkor: 21 év). Egy budapesti egyetem magyar nyelv és irodalom szakos alapképzésének hallgatói voltak, kevés fonetikai ismerettel rendelkeztek a teszt elvégzésének idején. A 44 felnőtt között 33 nő és 11 férfi szerepelt. Minden adatközlőtől (N=76) a 61 stimulusra kaptunk egy „rövid” vagy „hosszú” (bináris: 0 vagy 1) választ, így összesen 4636 választ elemeztünk.

Az elemzés során összesítettük az egyes stimulusokra adott válaszok arányát. Meghatároztuk a „hosszú” válaszok százalékos arányát a zárszakasz-időtartam függvényében. Eszerint például ha egyetlen válaszadótól sem kaptunk „hosszú” választ, akkor az arány 0%; amennyiben viszont minden válaszadó hosszúnak ítélte meg az adott hangot, az arány 100%-os. Ahol 50%-os az arány, ott a válaszadók fele-fele arányban ítélték a hangot szingletonnak, illetve geminátának (vagyis ennél a pontnál egyenlő a valószínűsége, hogy „rövid” vagy „hosszú” az elhangzott hang a percepció szerint). A kapott válaszokra logisztikus függvényt illesztettünk, és ábráztuk a válaszgörbét. A statisztikai elemzéshez bináris logisztikus regressziót alkalmaztunk az SPSS 20-as verziószámú szoftverben. Általánosított lineáris kevert modellt (GLMM) építettünk az adatokra, amelyben a függő változó a 0:1-es 'válaszok' voltak, a 'záridőtartamot' fix hatásként, az 'adatközlőket' random hatásként tüntettük fel. Továbbá elemeztük az 'életkor' (gyermek, felnőtt) és a 'képzéshely' (labiális, alveoláris, veláris) eredményekre gyakorolt hatását is (a 'záridőtartam' és az 'életkor', valamint a 'záridőtartam' és a 'képzéshely' interakciója is szerepelt a modellünkben mint fix hatás). Összehasonlítottuk továbbá a gyermekek csoportján belül a 7 éves és a 8 évesek eredményeit, valamint az eredetileg szingletonból és az eredetileg geminátából létrehozott stimulusokra adott válaszokat; majd nem parametrikus próbákkal (Wilcoxon, Friedman-teszt) kiértékeljük a különbségeket.

3. Eredmények

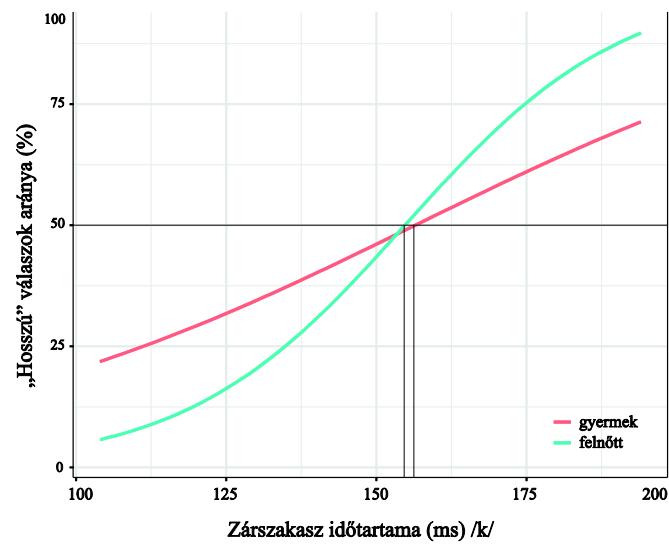
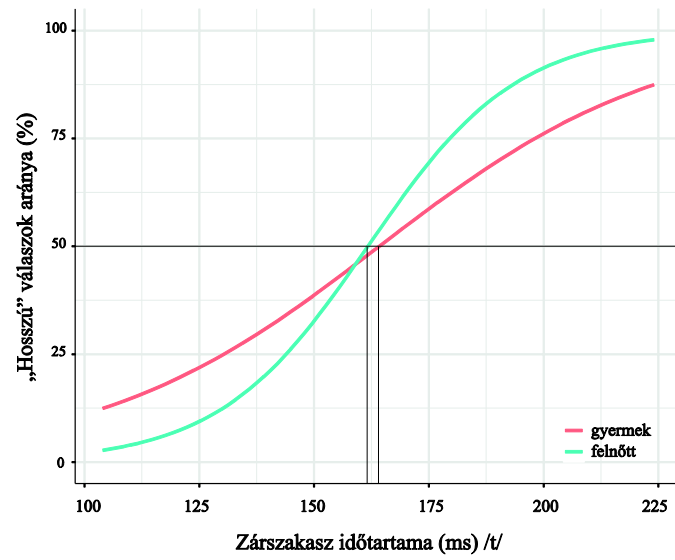
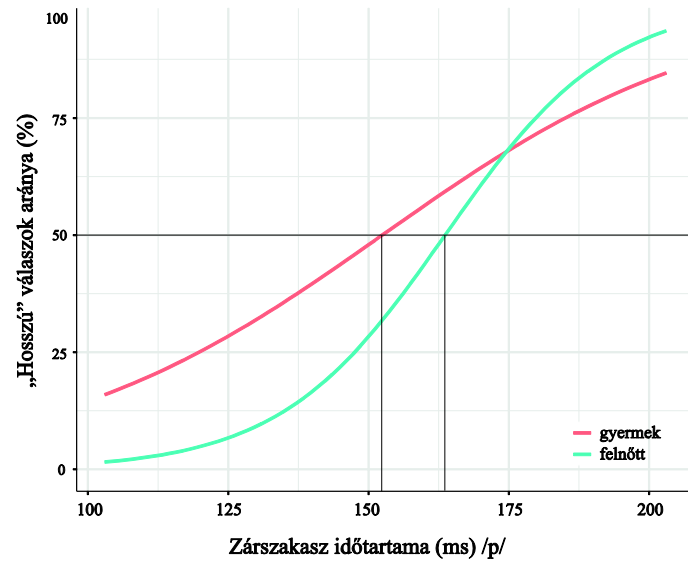
3.1 A mássalhangzó-hosszúság észlelése gyermekeknél és felnőtteknél

Az elemzés első lépéseként megvizsgáltuk a hallgatók ítéleteit az elhangzott mintában található expozívák zárszakaszának időtartamértékei alapján. Mivel a korábbi produkciós eredmények (Ridouane, 2010; magyarra: Neuberger, 2015) azt mutatják, hogy a záridőtartam a legfőbb akusztikai attribútum a fonológiai hosszúság megkülönböztetésében, azt vártuk, hogy a hallgatók is felhasználják ezt a paramétert a percepcióban. A hallgatói ítéletek alapján kialakult válaszgörbék (5. ábra) azt mutatják, hogy mind a felnőttek, mind a gyermekek döntéseit meghatározta a záridőtartam mindegyik expozíva esetén: relatíve rövid zárszakaszhoz „rövid” ítéletek, hosszú időtartamhoz pedig „hosszú” ítéletek adódtak nagy

arányban. A statisztikai elemzés kimutatta, hogy a záridőtartam szignifikánsan befolyásolja a hallgatók döntését: $F(1, 4629) = 869,142, p < 0,001$.

A felnőttek észlelésében a kísérleti anyagban szereplő minimum (100 ms körüli) záridőtartam esetén elvértve jelentek meg „hosszú” válaszok, míg a maximum (>200 ms) záridőtartamoknál a hallgatók döntése szinte egyöntetű volt a „hosszú” válaszokban. A felnőtt hallgatók egyöntetűen rövidnek ítélték a geminátákból 100 ms körüli záridőtartamúra csökkentett [p] és [t] hangokat, ám a [k] esetében 6,8%-uk „hosszú” választ adott a 100 ms körüli (legrövidebb záridőtartamú) explozívára. A 200 ms körüli explozívákra 82–97%-os arányban „gemináta” ítéletek születtek (itt is a [k] kategóriába sorolása mutatkozott a legbizonytalanabbnak). A felnőttek esetében a szigmoid (S-alakú) függvényt megközelítő eredmények a mássalhangzó-hosszúság észlelésének kategorikus voltára utalnak. A gyermekek válaszait tekintve a felnőttektől eltérő eredményt kaptunk. Egyfelől nem találtunk náluk 0% vagy 100%-os egyöntetű eredményt egyik hangnál sem, még a minimum és a maximum záridőtartam esetében sem. Másfelől az adatokra illesztett függvény alakja nem S-alakúnak mutatkozott, hanem megközelítőleg lineárisnak.

A statisztikai elemzést szerint annak valószínűségét, hogy a hallgató rövidnek vagy éppen hosszúnak ítélt meg egy adott időtartamú zárhangot a zárszakasz hosszán túl az adott hang képzéshelye is befolyásolja: $F(2, 4629) = 4,160, p = 0,016$. Azt is kimutatta az elemzés, hogy a záridőtartam a képzési hely szerint nem minden életkori csoportban van azonosan hatással a döntésekre: $F(5, 4629) = 7,964, p < 0,001$. Az adatokra illesztett görbékből jól látható, hogy labiális hang esetén található a legnagyobb különbség a gyermekek és a felnőttek válaszadása között (tekintve az 50%-nál meghúzható „percepciós határhoz” tartozó záridőtartamok különbségét a gyermekek és a felnőttek között).



5. ábra: Válaszgörbék és percepció határok gyermekeknél és felnőtteknél

A szingletonok és a gemináták közötti percepciók határt a nemzetközi szakirodalomban bevett gyakorlat szerint annál a záridőtartamnál húztuk meg, ami az adatokra illesztett logisztikus görbén az 50%-os válaszadáshoz kapcsolható (ahol a válaszadók fele rövidnek, fele hosszúnak ítélte meg az adott hangot). Ezt úgy határozhatjuk meg, hogy megvizsgáljuk azt a pontot, ahol az adatokra illesztett logisztikus görbe az 50%-os válaszadás egyenesét metszi. Az explozívák képzéshely szerinti különválasztása a percepciók határ kérdésében azért is indokolt, mivel a produkciós adatok is alátámasztják, hogy a zárszakasz időtartamát befolyásolja a zárhang képzési helye. Magyar nyelven Gósy & Ringen (2009) azt találta, hogy az átlagos záridőtartam a képzéshely hátrébb tolódásával csökkenő tendenciát mutat az intervokális zöngés explozíváknál izolált szavakban. Grácsi (2013) álszavakban mérte a zöngétlen zárhangok átlagos záridőtartamát, és szintén az előbb említett trendet támasztotta alá: 90 (± 13) ms a labiális, 72 (± 15) ms az alveoláris, valamint 70 (± 13) ms a veláris explozíváknál. A rövid zöngétlen explozívák esetében spontán beszédanyagban is megjelent a tendencia: 79 (± 11) ms a labiális, 71 (± 18) ms az alveoláris, 63 (± 18) ms a veláris mássalhangzók esetében (Neuberger 2015). Ugyanebben a kutatásban a gemináták átlagos záridőtartama pedig: 115 (± 20) ms-os, 122 (± 31) ms-os és 106 (± 27) ms-os átlagértékeket mutatott a képzéshellyel hátrafelé haladva. A percepciók határt illetően is hasonló különbségeket vártunk a hangok között a képzéshely függvényében. A percepciók eredmények a produkcióban megfigyelhető tendenciát mutatták a felnőtteknél: a képzési hely hátrébb tolódásával párhuzamosan a rövid-hosszú határ csökkenést mutatott (2. táblázat). A gyermekek percepciók határa azonban a labiális explozívánál rövidebb záridőtartamnál jelent meg, mint a másik két képzéshely esetén.

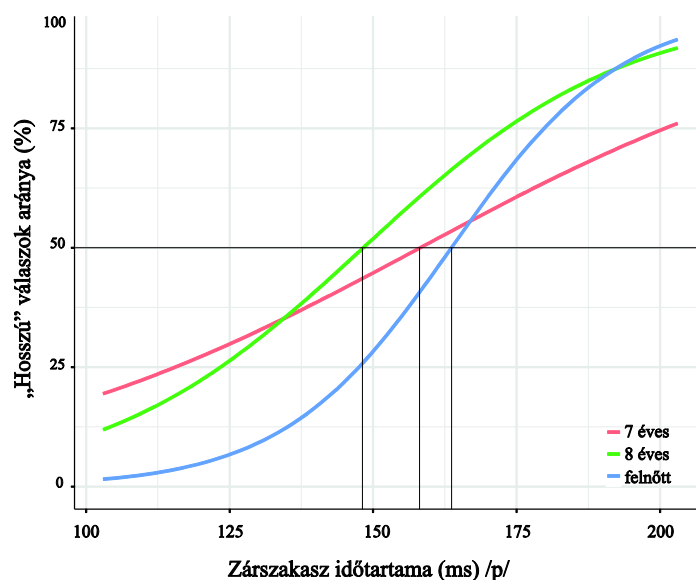
Kimutatható továbbá az is, hogy a gyermekek és a felnőttek észlelési határa az alveoláris, valamint a veláris explozíva esetében közel azonos záridőtartamoknál jelenik meg. Ezzel szemben a labiális zárhangnál a rövid és a hosszú kategória közötti határérték a gyermekek észlelésében 11 ms-mal alacsonyabb, mint a felnőttekében. Ez azt jelenti, hogy a gyermekek a 152 ms-nál rövidebb [p] hangot nagyobb valószínűséggel ítélik rövidnek, mint hosszúnak, viszont az ennél hosszabbakat potenciálisan hosszúnak. A felnőttek pedig 163 ms-os zárszakaszig valószínűsíthetően rövidnek ítélik meg a [p] hangot.

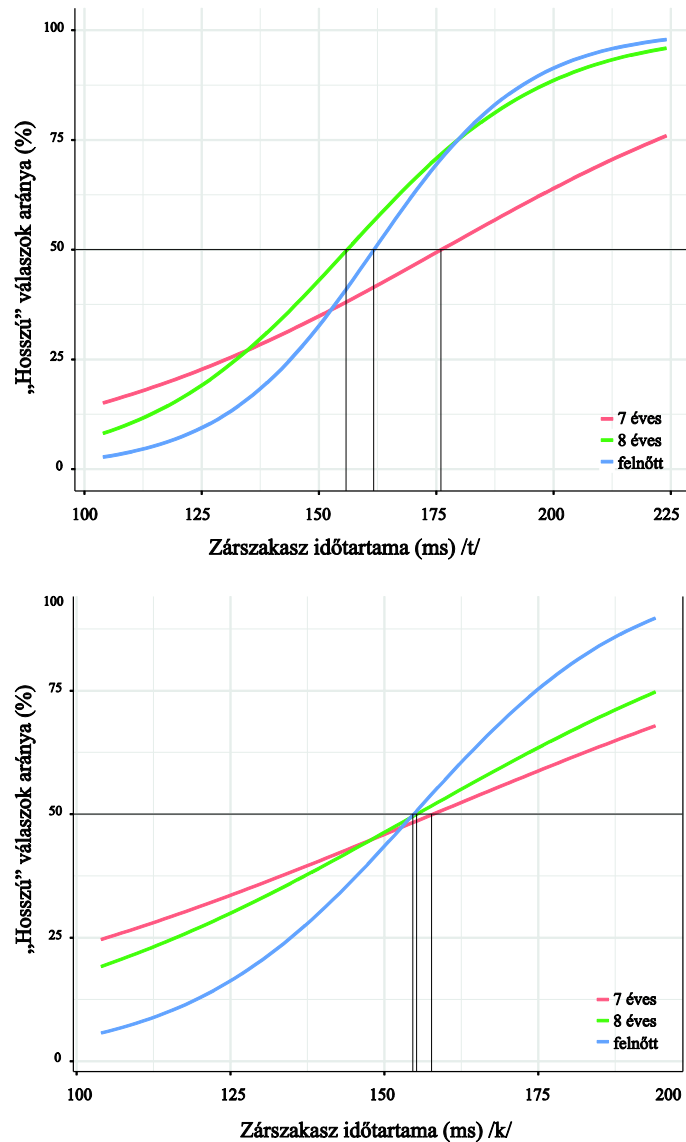
2. táblázat: Záridőtartamok a rövid-hosszú percepciók határnál a gyermekek és a felnőttek észlelésében

	[p]	[t]	[k]
Gyermekek	152 ms	164 ms	156 ms
Felnőttek	163 ms	161 ms	154 ms

3.2. Beszédszlelési különbségek a gyermekek életkora tekintetében

Az elemzés következő szakaszában különválasztottuk a 7 évesek és a 8 évesek eredményeit a korábbi szakirodalmi tapasztalatok okán, melyek szerint ez az az életkor, amely során jelentős változás tapasztalható az anyanyelvi fejlődés bizonyos területein, többek között a beszédhang-differenciálásban (lásd Bevezetés). Az így kialakított három csoport eredményeit a 6. ábrán szemléltetjük. Az adatokra épített általánosított lineáris kevert modell megerősítette, hogy az életkor befolyásolja a hallgatók válaszadását a záridőtartamok szerint: $F(2, 4624) = 16,774$, $p < 0,001$. A vizsgált életkorok közül a 7 évesek eredményei tértek el szignifikánsan a felnőttek eredményeitől ($p = 0,018$), de a gyermekek két csoportja, valamint a 8 évesek és a felnőttek eredményei között nem találtunk szignifikáns eltérést. Az adott hang képzéshelye függvényében szintén eltérő eredményeket kaptunk a különböző életkorú hallgatók döntéseit illetően: $F(6, 4624) = 4,283$, $p < 0,001$. A veláris hang esetén találtuk a legeggyöntetűbb válaszokat: a logisztikus függvények lefutása mind a hét- és a nyolcévesek, mind a felnőttek esetében hasonló képet mutat. A statisztikai elemzés a labiális és az alveoláris exploziva esetén szignifikáns különbséget mutatott az életkori csoportok között a „hosszú” válaszok százalékos arányát tekintve (Friedman $\chi^2 = 8,553$, $df = 2$, $p = 0,014$; illetve $\chi^2 = 8,435$, $df = 2$, $p = 0,015$), a veláris exploziva esetén azonban nem.





6. ábra: Válaszgörbék és percepció határok életkoronként

A rövid és a hosszú hangok közötti határ változó záridőtartam-értékeknél jelent meg a különböző életkorú hallgatóknál. Korábban láthattuk már, hogy a felnőtteknél a képzéshely szerint hátrafelé haladva egyre rövidebb időtartamokat kaptunk. A gyermekek egyik csoportjában sem jelent meg a tendencia. Két szélsőséges értéket találtunk az eredményekben: a hétévesek percepció határa a [t] hang esetében hosszabb, a nyolcévesek percepció határa pedig a [p] hang esetében rövidebb volt, mint a másik két hangnál vagy másik két csoportnál mért értékek.

3. táblázat: Záridőtartamok a rövid-hosszú percepció határnál életkoronként

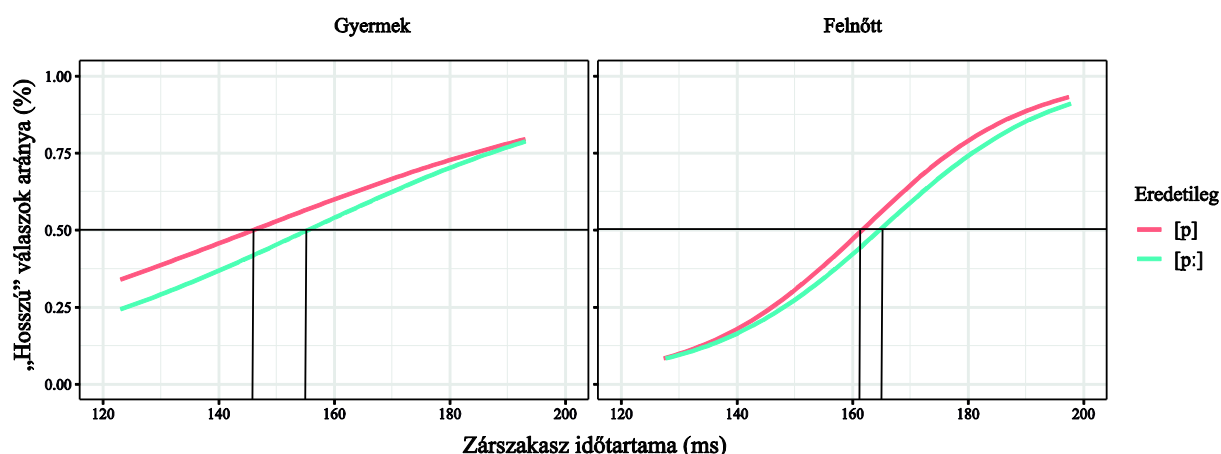
	[p]	[t]	[k]
Hétévesek	157 ms	175 ms	157 ms
Nyelcévesek	148 ms	154 ms	154 ms

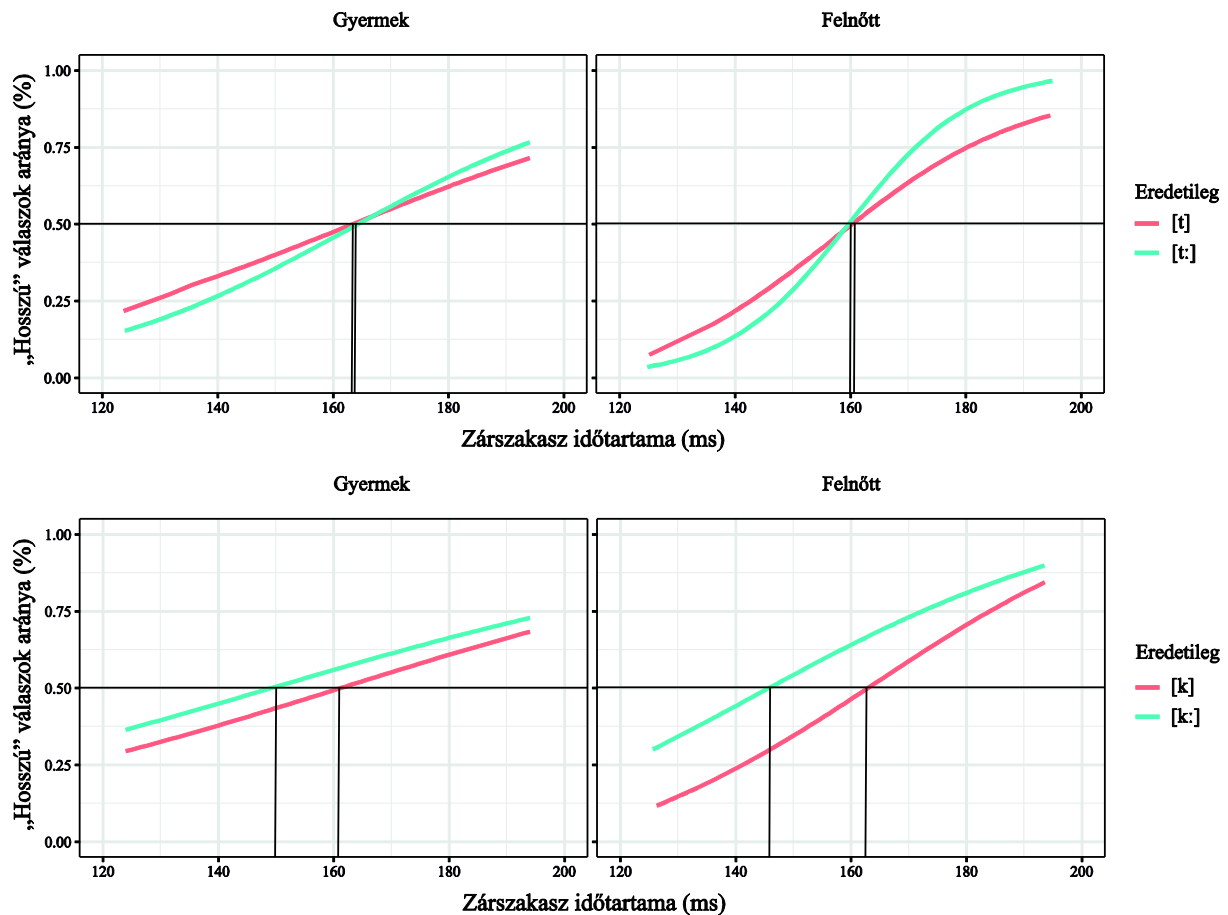
Felnőttek	163 ms	161 ms	154 ms
-----------	--------	--------	--------

3.3 Másodlagos felismerési kulcsok az észlelésben

Az elemzés utolsó lépésében megvizsgáltuk, hogy az a tényező, hogy az időtartam-manipuláció előtt az adott zárhang eredetileg szingleton vagy eredetileg gemináta volt, miként befolyásolja a hallgatók ítéleteit. Vajon a zárképzés időtartamának változtatása után marad-e valamilyen akusztikai jellemző, amely az eredeti hang hosszúsági jegyét magában hordozza, és amit a hallgatók felismernek az észlelés során? Ezt a kérdést azért érdemes körüljárni, mert a korábbi szakirodalmi eredmények (pl. Muller, 2003; Sato et al., 2012) felvetették a záridőtartam mellett további, másodlagos felismerési kulcsok lehetőségét az észlelésben. Ennek érdekében összesítettük a gyermek és felnőtt hallgatók válaszait különválasztva az eredetileg szingleton hangot tartalmazó hangszekvenciákat (iCi) az eredetileg geminátát tartalmazóktól (iC:i). A 0:1-es adatokra illesztett válaszgörbákat és a percepció határt ismét a képzési hely szerint vizsgáltuk a gyermekeknél és a felnőtteknél (7. ábra).

Az eredetileg szingletonból manipulált beszédhangokra adott válaszok csupán kis mértékben különböztek az eredetileg geminátákból alkotott hangokétól a válaszgörbék alakját és elhelyezkedését megfigyelve. A statisztikai elemzés szerint a felnőtteknél a veláris hang esetében tértek el a válaszok szignifikánsak az eredeti kvantitás szerint (Wilcoxon $Z = -2,524$; $p = 0,012$). A gyermekek válaszaiban egyik hangnál sem volt kimutatható szignifikáns különbség.





7. ábra: Válaszgörbék és percepció határok a hang eredeti kvantitása szerint

A rövid-hosszú percepció határt hasonló záridőtartamoknál húzhatjuk meg az eredetileg szingeton és az eredetileg gemináta stimulus esetén is (4. táblázat). A gyermekeknél a [p] és [k] hangnál 10 ms-os eltérést, míg a [t] hangnál semmilyen eltérést nem találtunk a válaszgörbék 50%-os értékei között az eredeti kvantitás függvényében. A felnőtteknél a [p] és a [t] hangnál 3 ms, illetve 1 ms-os különbséget találtunk. A legnagyobb eltérés a felnőttek [k] hangra adott válaszaiban adatolható (15 ms-os eltérés): a geminátából képzett rövid záridőtartamú stimulusokra a hallgatók nagyobb arányban adtak „hosszú” választ, mint az ugyanilyen zárszakaszú szingletonból képzett stimulusokra. Ez azt sugallja, hogy a hallgatók még rövidebb záridőtartam esetén is hajlamosabbak voltak hosszúnak ítélni a hangot, amely a manipuláció előtt gemináta volt. A percepció határ a szingletonból alkotott stimulusoknál 162 ms, míg a geminátából létrehozott stimulusoknál 147 ms. Ez azt jelenti, hogy előbbi esetben 162 ms-tól ítélték inkább hosszúnak a hangot, míg az eredetileg gemináta hangszekvenciákat már 147 ms-tól nagyobb arányban tartották hosszúnak, mint rövidnek.

4. táblázat: Záridőtartamok a rövid-hosszú percepció határnál az eredeti hang kvantitása szerint (er. = eredetileg, szin. = szingleton, gem. = gemináta)

	[p]		[t]		[k]	
	Er. szin.	Er. gem.	Er. szin.	Er. gem.	Er. szin.	Er. gem.
Gyermekek	146	156	164	164	161	151
Felnőttek	162	165	161	160	162	147

Összesítve az eredményeket (képzéshelytől és eredeti kvantitástól függetlenül), a rövid és a hosszú zöngétlen zárhangok közötti percepció határ a gyermekek észlelésében a zárszakasznak a 146–164 ms-os tartományában, a felnőtteknél pedig a 147–165 ms-os záridőtartamoknál található a jelen kísérlet anyagán elemezve.

4. Következtetések

A jelen tanulmányban a mássalhangzó-hosszúság észlelését vizsgáltuk a zöngétlen explozívák záridőtartamának függvényében. A beszédészlelés kutatásában régóta központi kérdés, hogy a hallgatók a beszédjel mely jellemzőit használják fel ahhoz, hogy fonetikai ítéleteket hozzanak. A korábbi nemzetközi és hazai kutatási eredmények alapján feltételeztük, hogy a hallgatók kvantitásra vonatkozó ítéleteihez elégséges kulcs a zárképzés időtartama. Eredményeink megerősítették ezt a hipotézist a vizsgálatban részt vett gyermekeknél és felnőtteknél is. Mindkét életkori csoportban szignifikánsan befolyásolta a záridőtartam az explozívákra adott „rövid” vagy „hosszú” válaszokat a hallott hangszekvenciákban. A 100 ms és 200 ms közötti tartományban megfigyelhető volt, ahogyan a hallgatók észlelése a „rövid” kategóriából átvált a „hosszú” kategóriába mindhárom explozíva esetén. Jóllehet a hallgatók döntése nem minden esetben mutatkozott egyöntetűnek még a minimum és maximum záridőtartamok esetén sem.

A kutatás második és a harmadik hipotézise a gyermekek és a felnőttek észlelése közötti eltérésekre vonatkozott. Az adatokra illesztett válaszgörbék alakjában megmutatkozó különbségek, valamint a statisztikai eredmények is megerősítették ezeket a feltevéseket. A felnőtteknél megjelent S-alakú görbék azt jelezték, hogy a záridőtartam kontinuum mentén van egy időintervallum, amikor a hallgatók ítéletei átváltanak szingletonból geminátába, míg a relatíve rövid vagy hosszú időtartamok (főként a skála két végén található minimum és maximum értékek) egyöntetűbb döntést eredményeznek. Az akusztikum és a percepció közötti kapcsolat nem lineáris volta (Stevens, 1989) a jelen kutatás felnőttekre vonatkozó eredményeiben is jól megmutatkozott. A gyermekek válaszaira illesztett függvényekben azonban nem rajzolódott ki a tipikus S-alakú görbe, az ő eredményeik sokkal inkább lineáris elrendezést mutattak. Még a vizsgálatunkban alkalmazott időtartamskála minimum és

maximum értékeinél sem született egyöntetű döntés a hosszúság megítélésében. A kvantitás kategoriális észlelése ezek szerint fejlődés eredménye, és a kategóriák megítélésében a 7-8 éves kornál későbbi életkorban válik a felnőttekéhez hasonló szintűvé a magyar anyanyelvű gyermekeknél.

A szingletonok és a gemináták közötti határhoz kapcsolódó időtartamértékeket már több nemzetközi kutatás vizsgálta mind a produkciós, mind a percepciós oldalról (pl. Esposito & Benedetto, 1999; Amano & Hirata, 2010, 2015). Tudomásunk szerint nem született még olyan tanulmány, amely a gyermekek észlelésében vizsgálta a kategóriák közötti választóvonalat. Vizsgálatunk fényt derített arra, hogy – bár a rövid és hosszú zárhangok észlelésében különbség mutatható ki gyermekek és felnőttek között –, a kategóriák észlelési határa nem feltétlenül tér el a felnőtteknél megjelenő időtartamértékektől. A különbségek akkor váltak szembetűnővé a kísérletünkben, amikor a 7 és a 8 évesek eredményeit különválasztva elemeztük, de ekkor is maximum 15 ms-os különbség volt adatolható a zárszakaszok időtartamában a felnőttekhez képest. Mindent összevetve a rövid és hosszú kategóriák közötti percepciós határ anyagunkban a 146 és 175 ms közötti tartományban nyilvánult meg: a labiális explozívánál 146–165 ms között, az alveoláris explozívánál 154–175 ms között, a veláris explozívánál pedig 147–162 ms között. Ezek az értékek összhangban vannak a nemzetközi kutatásokban talált percepciós határvonal értékeivel (a zöngétlen explozívák záridőtartamára: 120–180 ms, pl. Esposito & Benedetto, 1999; Yoshida et al., 2015; Zhou et al., 2021).

A felnőtteknél a veláris explozíva esetében különbséget találtunk aszerint, hogy eredetileg szingleton vagy eredetileg gemináta szerepelt a hangsorban. Ebből arra következtethetünk, hogy a zárszakasz időtartamán túl más akusztikai jellemzők is szerepet játszhatnak a kvantitás jegyének felismerésében, amelyek a három vizsgált hang közül a [k] hang esetében bizonyult szálisnak. A hosszúsági oppozíció másodlagos akusztikai attribútumának tekinthető például a felpattanás intenzitása, amelyben a produkciós vizsgálatok alapján különbség figyelhető meg a rövid és hosszú felpattanó zárhangok között (Neuberger & Beke, 2017). Továbbá a szingletonokat és a geminátákat megelőző magánhangzók időtartamának produkcióban megmutatkozó különbsége alapján feltételezhetjük, hogy ez a paraméter is hozzájárul a kontraszt pontos észleléséhez (Deme et al. 2019). Az eredetileg szingleton és gemináta hangok válaszgörbéinek ilyenfajta „elcsúszását”, azaz a percepciós határok különbségét találták más nyelvekben is. A [p] explozíva esetén mind a japán, mind a finn hallgatók válaszaiban megjelent az, hogy a szingleton/gemináta határ alacsonyabb záridőtartamhoz kapcsolódik abban az esetben, ha a manipuláció előtt a hang gemináta volt (Yoshida et al.,

2015). A [t] explozíva esetén az eredeti geminátákból képzett stimulusokban szintén kisebb záridőtartamokhoz (167–168 ms körül) kapcsolódott az 50%-os válaszadás, mint az eredetileg szingletonoknál (170–175 ms körül) (Zhou et al., 2021).

A gyermekeknél egyik hangnál sem volt kimutatható a válaszokban különbség aszerint, hogy a stimulus szingleton avagy gemináta volt az időtartam-manipuláció előtt. Ennek egy magyarázata lehet az, hogy az ő percepciók működésük ebben az életkor még kevésbé érzékeny ezekre a másodlagos felismerési kulcsokra, illetőleg hogy más felismerési kulcsokra támaszkodnak az egyes fonológiai megkülönböztető jegyek felismerésekor, mint a felnőttek. Szakirodalmi adatok alátámasztják, hogy a kisgyermekek a hallás-alapú információkat nagyobb (szótag méretű) egységekben dolgozzák fel, az idősebb gyermekek és a felnőttek pedig kisebb egységekre (fonémákra) támaszkodnak. Bekövetkezik tehát egy fejlődésbeli váltás, amelynek során a gyermekek már másképp súlyozzák be az észlelés alapegységeit, ahogyan a növekvő nyelvi tapasztalattal párhuzamosan a percepciók stratégiáik is változnak (Nitttrouer & Lowenstein, 2009).

A mindennapi életben az egyes beszédhangok fonémaosztályokba sorolása ritkán korlátozódik csupán egyetlen beszédhangra, általában nagyobb egységekben történik, olykor egy teljes szó azonosítását követően jön létre. Ebben az esetben a hosszúsági jegy megítélését az akusztikai jellemzőkön túl a szó szemantikai tartalma is támogatja. A fonémadöntésre számos kontextuális tényező hatással van: ilyen a beszédhang helye a szótagban, a szótagszerkezet vagy a szomszédos beszédhangok jellemzői, amelyek viszonylatában az adott hang akusztikai sajátosságai (többek között az időtartamértékei is) változatos képet mutathatnak. Továbbá érdemes megemlíteni, hogy a hosszúsági kontraszt funkcionális terheltsége viszonylag alacsony a magyar mássalhangzórendszerben; az oppozícióban álló minimálpárok egyik vagy mindkét tagja leggyakrabban levezetett (vagyis nem mögöttes, szótagban is szereplő) geminátát tartalmaz, avagy hangutánzó szó, jövevényszó vagy tulajdonnév. Mindez magyarázatul szolgálhat ahhoz, hogy miért okoz nehézséget az izolált beszédhangok differenciálása a gyermekeknek (és olykor a felnőtteknek is).

Továbbá a fonológiai fejlődésben meghatározó az adott nyelvben a hangelőfordulások gyakorisága; a gyakran hallott kategóriák közötti kontrasztok előbb épülnek be a nyelvspecifikus elsajátítás során, mint a ritkábban hallottak (Anderson et al., 2003). Mivel a magyarban a kvantitásoppozíciónak a mássalhangzók körében alacsony a funkcionális terheltsége, disztinkatív szerepe főként levezetett (és nem lexikális, morfémán belüli) alakokban, valamint hangutánzó, indulatszó, jövevényszó vagy tulajdonnév tagú minimálpárokból előforduló szembenállásokra korlátozódik, ezért a kihasználtság, illetve a

gyakoriság szempontjából periférikus jelenségről van szó. Ezzel állhat összefüggésben az, hogy a magyar gyermekek relatíve későn válnak képessé a mássalhangzókat hosszúság szerint fonématudatosan elkülöníteni a percepcióban. Holott az anyanyelvi oktatás-nevelés, az írás- és olvasástanítás során szükséges lenne, hogy a gyermekek beszédészlelése, fonológiai tudatossága elérjen egy olyan szintet, amely a sikeres tanulás alapjául szolgálhat.

Lényeges hangsúlyozni, hogy a jelen kísérletben kapott percepciós határok csak bizonyos megszorításokkal hasonlíthatók össze más vizsgálati eredményekkel. Eredményeink korlátozottsága részben a kutatáshoz felhasznált kísérleti anyagból fakad. Ez egyetlen női beszélő felolvasott hangszekvenciáin alapul, amelyben a magánhangzó spektrális-temporális jegyeit változatlanul hagytuk. Ebből kifolyólag egyrészt az artikulációs tempó szerinti különbségekre, illetve ezeknek a percepcióra gyakorolt hatására itt nem derülhetett fény. Másrészt, ismeretes, hogy a felpattanó zárhangok észlelése kontextus-függő, a mássalhangzó és a környező szegmensek relációja további befolyásoló tényező lehet az kategóriák azonosításában. Jelenlegi anyagunkhoz csupán az [i] magánhangzó-környezetben ejtett zöngétlen explozívákat választottuk. A továbbiakban mindenképpen fontosnak tartjuk megvizsgálni nem csupán az abszolút időtartamok hatását a kategóriák közötti különbségek felismerésében, hanem a singleton vs. gemináta mássalhangzók és a kontextus időviszonyainak szerepét az észlelésben – ahogyan azt a relációs akusztikai invariancián alapuló kutatások kiemelik. Mindez pontosabb képet adhat a statikus nyelvi kategóriák és a beszédprodukció, illetve beszédpercepció dinamikus működése közötti kapcsolatról.

Köszönetnyilvánítás

A kutatást a Bolyai János Kutatási Ösztöndíj és az MTA Posztdoktori Kutatói Pályázat támogatta.

Irodalom

Amano, S., & Hirata, Y. (2010). Perception and production boundaries between single and geminate stops in Japanese. *The Journal of the Acoustical Society of America*, 128(4), 2049–2058.

Amano, S., & Hirata, Y. (2015). Perception and production of singleton and geminate stops in Japanese: Implications for the theory of acoustic invariance. *Phonetica*, 72(1), 43–60.

Anderson, J. L., Morgan, J. L., & White, K. S. (2003). A statistical basis for speech sound discrimination. *Language and Speech*, 46, 155–182.

- Aoyama, K. (2000). *A psycholinguistic perspective on Finnish and Japanese prosody: Perception, production and child acquisition of consonantal quantity distinctions*. Boston/Dordrecht/London: Kluwer Academic Publishers.
- Blumstein, S. & Stevens, K. (1981). Phonetic features and acoustic invariance in speech, *Cognition*, 10, 25–32.
- Boersma, P. & Weenink, D. (2020) Praat: doing phonetics by computer [Computer program]. Version 6.1., retrieved 10 January 2020 from <http://www.praat.org>
- Deme, A., Bartók, M. Grácsi, T. E., Csapó, T. G., & Markó, A. (2019). Gemináták artikulációs szerveződése a magyarban. *Beszéd kutatás*, 27, 54–74.
- Esposito, A., & Di Benedetto, M. G. (1999). Acoustical and perceptual study of gemination in Italian stops. *The Journal of the Acoustical Society of America*, 106(4), 2051–2062.
- Fowler, C. (1986). An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics*, 14, 3–28.
- Gnanadesikan, A. (2004). Markedness and faithfulness constraints in child phonology. In R. Kager, J. Pater, and W. Zonneveld (eds.), *Constraints in phonological acquisition*, (pp. 73–108). Cambridge University Press.
- Gósy, M. (1995/2006). *GMP-diagnosztika*. Nikol, Budapest.
- Gósy, M. (2006): A beszédhangok megkülönböztetésének fejlődése. *Beszéd kutatás* 2006. 147–159.
- Gósy, M., & Horváth, V. (2006). Beszédfeldolgozási folyamatok összefüggései gyermekkorban. *Magyar Nyelvőr*, 2006, 470–481.
- Gósy, M. & Ringen, C. O. (2009). Everything you always wanted to know about VOT in Hungarian. In: *IXth International Conference on the Structure of Hungarian*, Debrecen, Hungary. http://icsh9.unideb.hu/pph/handout/Ringen_Gosy_handout.pdf
- Grácsi, T. E. (2013) Explozívák és affrikáták zöngésségének időviszonyai. *Beszéd kutatás*, 21, 94–120.
- Gyarmathy, D. & Horváth, V. (2010). A beszédhallás szerepe a beszédhang-differenciálásában. *Gyógypedagógiai Szemle*, 38(2), 126–135.
- Hankamer, J., Lahiri, A. & Koreman, J. (1989). Perception of consonant length: Voiceless stops in Turkish and Bengali. *Journal of Phonetics*, 17(4), 283–298.
- Hirata, Y., & Whiton, J. (2005). Effects of speaking rate on the single/geminate stop distinction in Japanese. *The Journal of the Acoustical Society of America*, 118(3), 1647–1660.
- Jordanidisz Á. (2015). Magyar anyanyelvű gyermekek fonológiai tudatosságának fejlődése 4 és 10 éves kor között. Doktori disszertáció. ELTE, Budapest.
- Khattab, G., & Al-Tamimi, J. (2013). Influence of geminate structure on early Arabic templatic patterns. In Vihman, M. M., & Keren-Portnoy, T. (Eds.). (2013). *The emergence of phonology: Whole-word approaches and cross-linguistic evidence*. (pp. 374–414). Cambridge University Press.
- Klatt, D. H. (1979). Speech perception: A model of acoustic–phonetic analysis and lexical access. *Journal of phonetics*, 7(3), 279–312.
- Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophic Transactions of the Royal Society B*, 369, 979–1000.

- Kunnari, S., Nakai, S., & Vihman, M. M. (2001). Cross-linguistic evidence for the acquisition of geminates. *Psychology of Language and Communication*, 5(2), 13–24.
- Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages* (Vol. 1012). Oxford: Blackwell.
- Lieberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of experimental psychology*, 54(5), 358–368.
- Lieberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21(1), 1–36.
- Lisker, L. (1958). The Tamil occlusives: short vs. long or voiced vs. voiceless. *Indian Linguistics, Turner Jubilee, 1*, 294–301.
- Muller, Jennifer S. (2003): The production and perception of word-initial geminates in Cypriot Greek, In *ICPhS-15*, 1867–1870.
- Nagy, J., Józsa, K., Vidákovich, T., & Tibor-Fazekasné Fenyvesi, M. (2004): *DIFER Programcsomag: Diagnosztikus fejlődésvizsgáló és kritériumorientált fejlesztő rendszer 4–8 évesek számára*. Mozaik Kiadó, Szeged.
- Neuberger, T. (2015). Durational correlates of singleton-geminate contrast in Hungarian voiceless stops. In *ICPhS 2015*. <http://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0422.pdf>
- Neuberger, T. (2016). Perception of consonant length opposition in Hungarian stop consonants. *The Phonetician*, 113, 6–23.
- Neuberger, T., & Beke, A. (2017) A zárfelpattanás spektrális jegyei a hosszúsági oppozíció függvényében. *Beszédkutatás*, 25, 7–23.
- Nittrouer, S., & Lowenstein, J. H. (2009). Does harmonicity explain children's cue weighting of fricative-vowel syllables?. *The Journal of the Acoustical Society of America*, 125(3), 1679–1692.
- Obrecht, D. H. (1965). Three experiments in the perception of geminate consonants in Arabic. *Language and Speech*, 8(1), 31–41.
- Olaszy, G. (2006). *Hangidőtartamok és időszerkezeti elemek a magyar beszédben*. Nyelvtudományi Értekezések 155. Budapest: Akadémiai Kiadó.
- Payne, E., Post, B., Garmann, N. G., & Simonsen, H. G. (2017). The acquisition of long consonants in Norwegian. In Kubozono, H. (Ed.). (2017). *The phonetics and phonology of geminate consonants* (Vol. 2). (pp. 130–162). Oxford University Press.
- Pickett, E. R., Blumstein, S. E. & Burton, M. W. (1999). Effects of speaking rate on the singleton/geminate consonant contrast in Italian. *Phonetica*, 56(3-4), 135–157.
- Ridouane, R. (2010). Geminates at the junction of phonetics and phonology. *Papers in laboratory phonology*, 10, 61–90.
- Sato, Y., Kato, M., & R. Mazuka, R. (2012). Development of single/geminate obstruent discrimination by Japanese infants: Early integration of durational and nondurational cues. *Developmental Psychology*, 48(1), 18–34.
- Siptár, P., & Grácz, T. E. (2014). Degemination in Hungarian: Phonology or phonetics? *Acta Linguistica Hungarica*, 61, 443–471.

- Slifka, J. (2006). Acoustic cues, landmarks, and distinctive features: a model of human speech processing. *ECTI Transactions on Computer and Information Technology (ECTI-CIT)*, 2(2), 91–96.
- Sonu, M., Kato, H., Tajima, K., Akahane-Yamada, R. & Sagisaka, Y. (2013). Non-native perception and learning of the phonemic length contrast in spoken Japanese: training Korean listeners using words with geminate and singleton phonemes. *Journal of East Asian Linguistics*, 22(4), 373–398.
- Stevens, K. N. (1989). On the quantal nature of speech. *Journal of phonetics*, 17(1-2), 3–45.
- Stevens, K. N. (2002). Toward a model for lexical access based on acoustic landmarks and distinctive features. *The Journal of the Acoustical Society of America*, 111(4), 1872–1891.
- Stevens, K. N. & Klatt, D. H. (1974). Role of formant transitions in the voiced-voiceless distinction for stops. *The Journal of the Acoustical Society of America*, 55(3), 653–659.
- Stevens, K. N., & Blumstein, S. E. (1978). Invariant cues for place of articulation in stop consonants. *The Journal of the Acoustical Society of America*, 64(5), 1358–1368.
- Sussman, H. M., Mccaffrey, H. A. & Matthews, S. A. (1991). An investigation of locus equations as a source of relational invariance for stop place categorization. *The Journal of the Acoustical Society of America*, 90(3), 1309–1325.
- Tsukada, K., Cox, F., Hajek, J., & Hirata, Y. (2018). Non-native Japanese learners' perception of consonant length in Japanese and Italian. *Second Language Research*, 34(2), 179–200.
- Vihman, M. M. and Velleman, S. L. (2000). The construction of a first phonology. *Phonetica*, 57, 255–66.
- Wilson, A., Kato, H. & Tajima, K. (2005). Native and non-native perception of phonemic length contrasts in Japanese: Effects of speaking rate and presentation context. *The Journal of the Acoustical Society of America*, 117(4), 2425–2425.
- Yoshida, K., de Jong, K. J., Kruschke, J. K., & Päiviö, P. M. (2015). Cross-language similarity and difference in quantity categorization of Finnish and Japanese. *Journal of Phonetics*, 50, 81–98.
- Zhou, Y., Nakamura, Y., Mugitani, R., & Watanabe, J. (2021). Influence of prior auditory and visual information on speech perception: Evidence from Japanese singleton and geminate words. *Acoustical Science and Technology*, 42(1), 36–45.